

Entrega Etapa Prototipação

Equipe 28 – Noble Eagles

Produto Final

Nosso produto final será uma api pública pela qual o SINE e outros stakeholders possam acessar a propensão, descrita como uma probabilidade, de um dado pretendente ser contratado para uma determinada vaga.

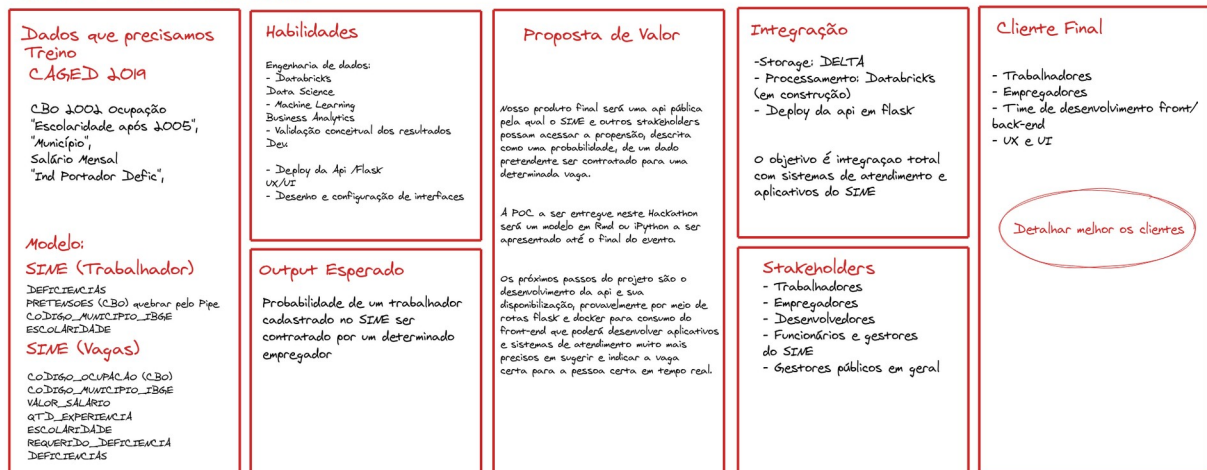
A POC a ser entregue neste Hackathon será um modelo em Rmd ou iPython a ser apresentado até o final do evento.

Os próximos passos do projeto são o desenvolvimento da api e sua disponibilização, provavelmente por meio de rotas flask e docker para consumo do front-end que poderá desenvolver aplicativos e sistemas de atendimento muito mais precisos em sugerir e indicar a vaga certa para a pessoa certa em tempo real.

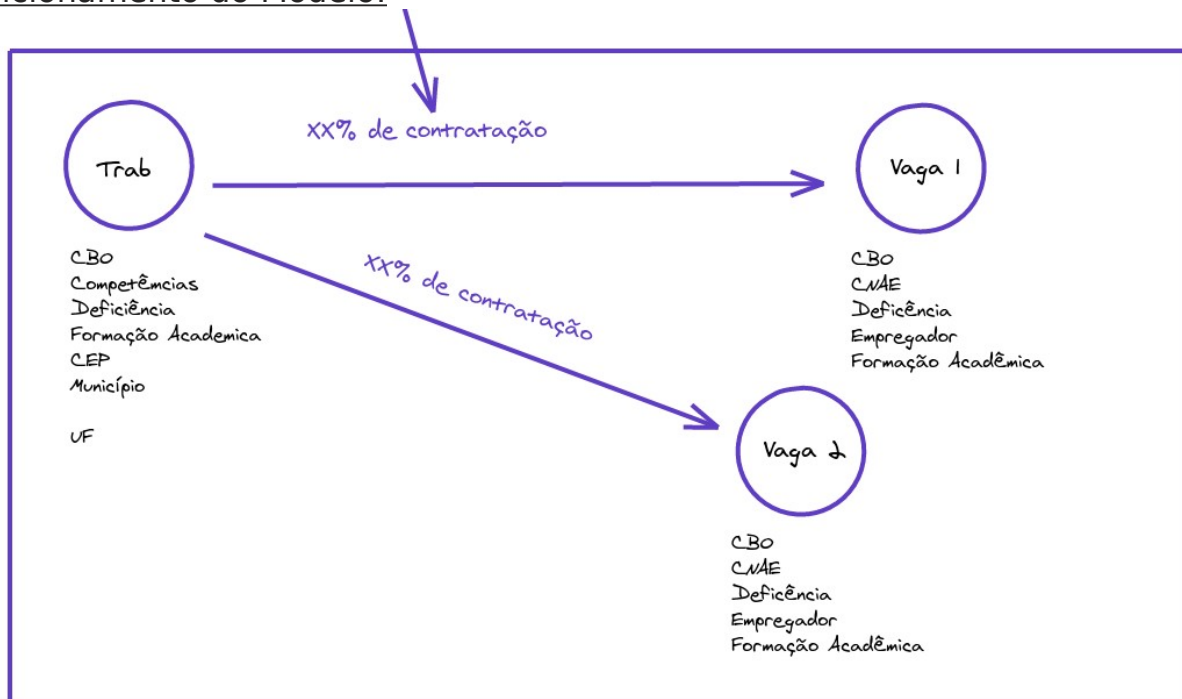
Detalhamento do desenvolvimento

CANVAS:

CANVAS Ciência de Dados



Funcionamento do Modelo:

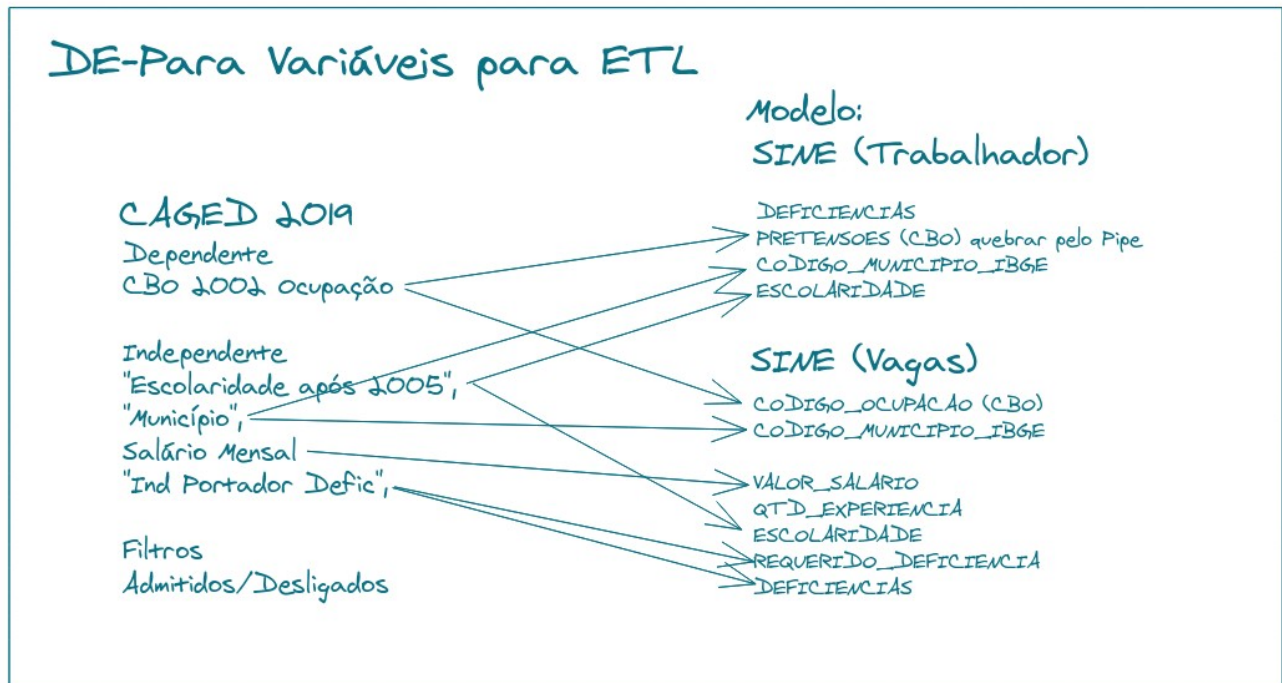


O algoritmo será construído com base nas informações da base do CAGED e será treinado com um grande volume de informações para que, com base nas informações também disponíveis na base do SINE possam alimentar o modelo em tempo real e responder a probabilidade de contratação para uma dada relação colaborado-vaga.

O modelo será treinado com base nos dados do Sudeste da base da CAGED, validado com base nos dados do Sul e aplicado aos dados disponibilizados pelo SINE.

Detalhamento do ETL

Para isto avaliamos os dicionários de dados disponíveis em cada base e planejamos como o ETL deveria ser realizado para assegurar que haja informações suficientes para a estimação de um MVP do modelo.



Primeiramente, tentamos uma abordagem tradicional com bancos relacionais do Azure, no entanto a performance foi insatisfatória e pivotamos a solução para o Databricks.

Desenho final do protótipo em desenvolvimento

Armazenamento de dados Delta com apoio de solução gratuita do Databricks.
MVP do modelo baseado em regressões logísticas
Apresentação em Rmd dos primeiros resultados